CLAIMS

What is claimed is:

1. In a computer system comprising a server machine and a client machine, a text-to-speech synthesis method comprising:

   a) obtaining a normalized text;

   b) selecting acoustic units corresponding to said normalized text from a database accessible to said server machine, said database storing a predetermined number of possible acoustic units;

   c) transmitting compressed acoustic units from said server machine to said client machine, wherein said compressed acoustic units are obtained by compressing said selected acoustic units using a compression method selected in dependence on said predetermined number of possible acoustic units; and

   d) in said client machine, concatenating said selected acoustic units.

2. The method of claim 1, further comprising generating prosody data corresponding to said normalized text and transmitting said prosody data from said server machine to said client machine, wherein step (d) comprises concatenating said selected acoustic units in dependence on said prosody data.

3. The method of claim 1 wherein step (d) further comprises concatenating said selected acoustic units with at least one cached acoustic unit, wherein said cached acoustic unit is cached on said client machine.

4. The method of claim 1, further comprising normalizing a standard text to obtain said normalized text.

5. The method of claim 1 wherein said possible acoustic units are compressed possible acoustic units, and wherein said compressed acoustic units are compressed before being stored in said database.

1
2
3

6.   The method of claim 1 wherein parameters of said compression method are selected to minimize the amount of data transmitted between said server machine and said client machine for each possible acoustic unit.

1
2
3

7.   The method of claim 6 wherein parameters of said compression method are further selected to achieve a minimum quality for each possible acoustic unit.

1
2

8.   The method of claim 1 wherein steps (c) and (d) are performed simultaneously for sequential acoustic units.

1
2
3
4
5
6
7
8

9.   In a server machine, a text-to-speech synthesis method comprising:

a)   obtaining a normalized text;

b)   selecting acoustic units corresponding to said normalized text from a database storing a predetermined number of possible acoustic units; and

c)   transmitting compressed acoustic units to a client machine, wherein said compressed acoustic units are obtained by compressing said selected acoustic units using a compression method selected in dependence on said predetermined number of possible acoustic units.

1
2
3

10.   The method of claim 9, further comprising generating prosody data corresponding to said normalized text and transmitting said prosody data to said client machine.

1
2

11.   The method of claim 9, further comprising normalizing a standard text to obtain said normalized text.

1
2
3

12.   The method of claim 9 wherein said possible acoustic units are compressed possible acoustic units, and wherein said compressed acoustic units are compressed before being stored in said database.

1
2
3

13.   The method of claim 9 wherein parameters of said compression method are selected to minimize the amount of data transmitted to said client machine for each possible acoustic unit.

14. The method of claim 13 wherein parameters of said compression method are further selected to achieve a minimum quality for each possible acoustic unit.

15. In a client machine, a text-to-speech synthesis method comprising:

a) receiving compressed acoustic units corresponding to a normalized text from a server machine, said compressed acoustic units being selected from a predetermined number of possible acoustic units and compressed using a compression method selected in dependence on said predetermined number of possible acoustic units;

b) decompressing said compressed acoustic units to obtain decompressed acoustic units; and

c) concatenating said decompressed acoustic units.

16. The method of claim 15, further comprising receiving prosody data corresponding to said normalized text from said server machine, wherein step (c) comprises concatenating said decompressed acoustic units in dependence on said prosody data.

17. The method of claim 15 wherein step (c) further comprises concatenating said decompressed acoustic units with at least one cached acoustic unit.

18. The method of claim 15 further comprising, before step (a), transmitting a standard text corresponding to said normalized text to said server machine.

19. The method of claim 15 further comprising, before step (a), normalizing a standard text to obtain a normalized text, and transmitting said normalized text to said server machine.

20. The method of claim 15 wherein parameters of said compression method are selected to minimize the amount of data transmitted to said client machine for each possible acoustic unit.

21. The method of claim 20 wherein parameters of said compression method are further selected to achieve a minimum quality for each possible acoustic unit.

22. The method of claim 15 wherein steps (a), (b), and (c) are performed simultaneously.

23. A text-to-speech synthesis system comprising:
   a) a database of predetermined acoustic units;
   b) a server machine in communication with said database for selecting ones of said acoustic units corresponding to a normalized text and for generating prosody data corresponding to said normalized text; and
   c) a client machine in communication with said server machine for concatenating said selected acoustic units in dependence on said prosody data;
   wherein said server machine transmits compressed acoustic units to said client machine, and wherein said compressed acoustic units are obtained by compressing said selected acoustic units using a compression method selected in dependence on said predetermined acoustic units.

24. The system of claim 23 wherein said client machine contains at least one cached acoustic unit.

25. The system of claim 23 wherein said server machine normalizes a standard text to obtain said normalized text.

26. The system of claim 23 wherein said client machine normalizes a standard text to obtain said normalized text and transmits said normalized text to said server machine.

27. The system of claim 23 wherein said predetermined acoustic units in said database are compressed predetermined acoustic units.

1
2
3

28.   The system of claim 20 wherein parameters of said compression method are selected to minimize the amount of data transmitted between said server machine and said client machine.

1
2
3

29.   The system of claim 28 wherein parameters of said compression method are further selected to achieve a minimum quality for each predetermined acoustic unit.

1
2
3
4
5
6
7
8
9
10

30.   A program storage device accessible by a server machine, tangibly embodying a program of instructions executable by said server machine to perform method steps for a text-to-speech synthesis method, said method steps comprising:

a)   obtaining a normalized text;

b)   selecting acoustic units corresponding to said normalized text from a database storing a predetermined number of possible acoustic units; and

c)   transmitting compressed acoustic units to a client machine, wherein said compressed acoustic units are obtained by compressing said selected acoustic units using a compression method selected in dependence on said predetermined number of possible acoustic units.

1
2
3

31.   The device of claim 30 wherein said method steps further comprise generating prosody data corresponding to said normalized text and transmitting said prosody data to said client machine.

1
2

32.   The device of claim 30 wherein said method steps further comprise normalizing a standard text to obtain said normalized text.

1
2
3

33.   The device of claim 30 wherein said possible acoustic units are compressed possible acoustic units, and wherein said compressed acoustic units are compressed before being stored in said database.

1
2
3

34.   The device of claim 30 wherein parameters of said compression method are selected to minimize the amount of data transmitted to said client machine for each possible acoustic unit.

1

2

3

35. The device of claim 34 wherein parameters of said compression method are further selected to achieve a minimum quality for each possible acoustic unit.